From: McCormack, Craig (ECY)
To: Bradley, Dave (ECY)
Cc: Kissinger.Lon@epamail.epa.gov; Hankins, Martha (ECY); McCormack, Craig (ECY)
Subject: FW: FW: Chapter 6 questions for Craig
Date: Monday, August 27, 2012 2:28:20 PM

Dave: my responses below and note Lon has responded to your first set of questions/thank you Lon for your input/Craig

1. **What is an outlier? (I think this is covered by the earlier information provided by Craig and Nyak)**

2. **Do the various statistical approaches used in the tribal studies fall in the range of scientifically defensible methods?**

Sechena et al, 2003-API Study @ page 258-259 - specifically noted treatment of outliers. "…all observed values greater than 3 standard deviations (SD) above the mean was substituted by a smaller value (mean plus 3SD) within each seafood group. After the treatment of outliers for each of the individual seafood categories, the "all seafood" consumption rate was computed as the sum of all individual seafood subcategories. Again, the outliers in the "all seafood" category were adjusted downward to a value of its mean + 3SD. Finally, a readjustment was carried out to reflect the fact that the overall "all seafood" rate was the sum of the individual seafood categories, proportionately allocated using the percentage of each subcategory in the "all seafood".

Toy et al., 1996 (Squaxin Island and Tulalip Tribes). Treatment of outliers. "...outliers were handled by recoding them to the largest reported consumption rate within three standard deviations of the arithmetic mean. This recoding was done separately by tribe and separately within each fish group subtotal: anadromous, pelagic, bottom, shellfish, and other fish. These corrections were applied separately for adults and for children."

Duncan, 2000 – Suquamish Tribal Survey. The authors of the Suquamish tribal survey did not consider statistical outliers as outliers of estimates of consumption. These high estimates were considered as actual estimates of consumption of subsistence populations within the Suquamish Tribal population of fish consumers. The survey notes the following @ page 70-71: "A number of high consumption rates were included in calculations of the mean, standard errors and percentiles, in contrast to some preceding survey (e.g., Toy et al.) where high values were considered as outliers and were truncated to a smaller value, such as the mean plus three standard deviations.

In the Suquamish Survey, these high values were believed to reflect actual high consumption and were not treated as outliers. In fact, the high values have no influence on the percentiles reported here for all seafood groupings (A-G) and all larger groups (all finfish, all shellfish, all seafood) with the single exception of "all finfish," where the 95th percentile would be slightly higher due to the inclusion of the high consumption rate reported by one respondent (4.570 g/kg/day) rather than the value that would have been used had it been truncated. Thus, percentiles are virtually unaffected by the use of these large consumption rates and calculations of percentiles.

It is possible that mean consumption rate may have been affected, though these changes would be small. For example, if the two highest consumption rates for all seafood reported by two individuals (18.4 and 14.8 g/kg/day) had been truncated to the mean plus three standard deviations (12.364), the revised consumption rate would have been 2.61 instead of 2.71, a minor difference. For Group G, if the one high rate reported (1.344 g/kg/day) had been truncated to the mean plus three standard deviations (0.78), the mean would have been revised to 0.45 g/kg/day rather than 0.52 g/kg/day, a 14% decrease."

1994 CRITFC Survey, @ page 26. "Outliers, those data points that seemed unreasonably high due to discontinuity in distribution, were identified in responses to some survey questions. A total of five outliers were identified and these data points were ignored in all calculations." This was the first of the tribal surveys conducted and the concept of subsistence populations was not considered during the data analysis. Other authors - Harper, Harris, and Donatuto – have noted that ignoring these estimates may not account for estimates of fish consumption that reflect subsistence populations.

- **Is there some standard guidance on this issue/rules of thumb/criteria for selecting an approach?**

  The EPA 1998 Guidance for Conducting Fish and Wildlife Consumption Surveys offers little in terms of statistical outlier – it notes only: "If variances are high, it is appropriate to examine the data for outliers and apply the appropriate nonparametric test." The 1992 Consumption Surveys for Fish and Shellfish, A review and Analysis of Survey Methods offers no guidance about outliers that I could locate with the word search function. From Statistics, Concepts and Controversies, 3rd Edition, by David S. Moore notes that a statistical outlier are data points or values that stand apart from the rest or observations that fall outside the pattern. (@page 196 & 256 A definition similar to the one used by CRITFC to ignore values.

  Given that the API study has been described by Faustman as the Gold Standard it would seem that 3+SD from the mean is a reasonable statistical approach to handle outliers. CRITFC ignored the outliers in the analysis noting that 5 outliers were different from the other data points – a definition similar to above but handled differently than the other surveys. It would seem that each survey provided a rationale for the statistical treatment of outliers and each used a defensible, at least within a range of technical defensibility.

- **How is this issue handled in other types of studies (is the Suquamish methodology outside the norm?)**

  There is variation across the surveys regarding statistical treatment of outliers. However, the CRITFC study should be viewed as the first of a series with the other studies evolving in terms of treatment of the outliers. The Suquamish was the first to consider the surveyed fish consuming populations having different fish consumers within the tribal fish consuming population -

- identified as subsistence.

3. **How much does the statistical method for outliers impact the study results (e.g. what if the Tulalip methods were used with the Suquamish data)?**

   See discussion with the Suquamish dataset

4. **What are the implications for establishing rates that reflect annual/chronic exposure? (interaction with short-term episodic study issue?)**

   The tribal populations are fish consumers with fish being the primary source of protein. Hence, fish are not episodically consumed but are routinely consumed with dietary recall bias minimized.

From: Lon Kissinger [mailto:Kissinger.Lon@epamail.epa.gov]
Sent: Monday, August 27, 2012 2:07 PM
To: McCormack, Craig (ECY)
Cc: Marcia Bailey; Marc Stifelman
Subject: Re: FW: Chapter 6 questions for Craig

Hi Craig,

Again, the problem is that the short recall period for a 24 hour survey creates a bimodal distribution of consumers and non consumers, regardless of the season when a 24 hour recall survey is administered. The question remains as to what to do with non consumers. Lorenz is correct in that the average of the Suquamish 24 recall data is less than the food frequency questionnaire. And...it should be, because the 24 hour recall results included 55% of respondents who consumed no seafood!!! The use of non consumers in computing FCR statistics is fundamentally unacceptable to Ecology and EPA. The staggering of 24 hour recall events throughout the year would likely make per capita seafood consumption estimates more reflective of overall annual consumption, but again, the fundamental structure of the survey still leads to a bimodal distribution of no-consumption vs. something ranging between one to two times a standard meal size. The Tooze adjustment is obviously important here. I wouldn't say however, that there is a systematic bias one way or the other for FFQ vs. 24 hour results.

Of significance however, is that for each individual 24 hour recall and FFQ results were positively correlated, indicating correspondence between the two approaches.

Lon Kissinger
Toxicologist
Office of Environmental Assessment, Risk Evaluation Unit
U.S. EPA - Region 10, Suite 900
Mail Stop: OEA-095
1200 6th Ave.
Seattle, WA 98101

kissinger.lon@epa.gov

206-553-2115 voice
206-553-0119 FAX

"McCormack, Craig (ECY)" ---08/27/2012 09:44:00 AM---Hi Lon: Dave et al is in the process of

completing Ecology's response to comments on the FC Rate TSD

From: "McCormack, Craig (ECY)" <cmcc461@ECY.WA.GOV>
To: Lon Kissinger/R10/USEPA/US@EPA
Date: 08/27/2012 09:44 AM
Subject: FW: Chapter 6 questions for Craig

Hi Lon:
Dave et al is in the process of completing Ecology's response to comments on the FC Rate TSD. To help respond, I have provided Dave with articles by Tooze et al 2006, Subar et al 2006, and Dodd et al 2006 as technical references regarding 24 hour dietary recall studies – used in combination the 24 hour recall and FFQ provide information over the long term and the ability to cross check the short term recall responses with long term dietary portion sizes. Dave's commentary seems focused on variability related to fish consumption estimates over the short term Vs over the long term which is a problem for national data where fish is consumed less frequently, episodically, and derived based on short term recall. Do you have any additional insights or references?

Regarding EPA Region – 10 framework; my understanding is that the framework complies with the EPA information guidelines-both in terms of the information hierarchy and quality of information in terms of a quality assurance program. Any insights?
Thanks/Craig

From: Bradley, Dave (ECY)
Sent: Monday, August 27, 2012 8:55 AM
To: McCormack, Craig (ECY)
Cc: Hankins, Martha (ECY)
Subject: Chapter 6 questions for Craig

Craig –

I am still wading through the response to comments. There are still a few responses where I need help.

I have attached to comments with some questions I have relative to crafting a response.

Could you take a look at those and provide any insights you can?

Thanks

Dave
[attachment "Chapter 6 questions for Craig.docx" deleted by Lon Kissinger/R10/USEPA/US]